Autonomous Spectrum Assignment of White Space Devices

Chaitali Diwan(✉), Srinath Srinivasa, and Bala Murali Krishna

International Institute of Information Technology, Bangalore, India
{chaitali.diwan,balamurali.krishna}@iiitb.org, sri@iiitb.ac.in

Abstract. White-space spectrum has temporal and spatial variations, and fragmentation, making the spectrum assignment for devices in this space challenging. In this paper, we propose an autonomous agent model for spectrum assignment of white space devices at a given location. Each white space device (WSD) acts autonomously out of self-interest, choosing a strategy from its bag of strategies. It obtains a payoff based on its choice and choices made by all other WSDs. Based on the payoffs received by different strategies, WSDs evolve their strategic profile over time. This has the effect of demographic changes in the population which is published as demographic profile by the Master. WSDs are expected to choose a strategy with a probability distribution based on this, for optimising network utilisation. In evaluation runs, network utilisation levels in such an approach are found to be high, and approaching optimal values computed in a centralised fashion.

Keywords: White spaces · Dynamic spectrum access
Multi-agent systems · Evolutionary game theory
White space database · Optimising spectrum utilisation

1 Introduction

White space refers to licensed radio spectrum that is not being utilised by the licensee. Across the world, white space contributes to a lot of available spectrum that is wasted. Recent research has enabled the use of this spectrum for broadband internet access using a paradigm called dynamic spectrum access. In this model, the primary owner and the licensee, (typically terrestrial TV) would retain primacy over the spectrum, while letting the secondary users or white space devices (WSD) utilise this spectrum whenever a primary user is not using it. WSDs have to ensure that none of their transmissions interfere with the primary users. The most efficient way to ensure this is to use *geo-location databases* which has now become a standard way for operating in white spaces in most of the countries. Geo-location white space spectrum database (WSDB) is an authoritative source that publishes the spectral and temporal availability of the free channels in a given location. It implements the rules of authority and uses the stored information of the licensed primary users, secondary users, terrain

P. Marques et al. (Eds.): CROWNCOM 2017, LNICST 228, pp. 37–48, 2018.
https://doi.org/10.1007/978-3-319-76207-4_4

information, tower specific parameters, interference, etc. It also considers re-use of a channel and enforces social and environmental norms, such as allowing only low power transmissions near a contour or a national border.

White space devices can be of different types with varying characteristics, and with the proliferation of Internet of Things (IoT), the diversity in WSDs is very high, making it a challenge to manage their use of the free spectrum as well as their interaction with the WSDB. For this reason, WSDs are broadly categorised into two types: *Master WSD* and *secondary WSDs*. In any given WSD deployment, the Master WSD communicates directly with the WSDB, while secondary WSDs always communicate with WSDB through the Master WSD [1,2]. (We refer to secondary WSD as "WSD" and Master WSD as "Master" through out the paper.) A new deployment is started by the Master registering itself with the WSDB, and subsequently, secondary WSDs registering with the Master. Whenever a WSD wants to operate in the white space spectrum, it sends a request to the Master which in-turn forwards the request to the WSDB, and obtains the list of available frequency ranges along with the maximum transmission power, start and end time of spectrum availability for each of the spectrum frequency ranges. It notifies the WSDB when it uses a channel. When a primary user needs the spectrum, WSD must cease to operate in that frequency range.

In the above protocol, the main challenge is that of dynamic spectrum assignment to cater to the disparate needs of the different WSDs. One way to approach this problem is for the Master to perform all the allocations. However, the Master is also usually a low power device, and cannot afford the costly computations needed for optimally allocating spectrum in a dynamic environment. In addition, in a centralised decision-making model, the rest of the system will need to cease operations until the Master completes its decision-making and performs channel allocations.

To address this problem, we propose an autonomous-agent model for spectrum assignment of white space devices at a given location. In this model, each WSD acts autonomously out of self-interest, to allocate spectrum for itself. Based on its choice and the choices made by all other WSDs, it obtains a payoff. The payoff acts as a feedback function for the WSD to refine its strategy for making its choice. Such a system is trained on a given workload profile, until such time that the distribution of strategies (also called the "demographic profile") stabilises across the population.

In a deployed system, the role of the Master is limited to profiling the load based on the traffic data and the desired demographic profile based on the training. The Master publishes both these data in order to enable WSDs to alter the probability with which they choose a given strategy.

1.1 Related Work

Initial efforts towards spectrum allocation in white space networks were based on variants of graph colouring algorithms, some examples include [3–5]. These algorithms considered fixed topology or topology with infrequent updates. Spectrum assignment problem considered in [6] also assumes fixed spectrum availability.

Cao and Zheng in [7], Nie and Comaniciu in [8] and Suris *et al.* in [9] studied the spectrum assignment problem in white space network as cooperative game which is useful in scenarios where a single service provider deploys large number of wireless devices and enforces collaboration agreements among them.

Centralised approaches to spectrum allocation have been investigated in [10–13]. Although they have very high spectrum utilisation, they are not very suitable for dynamic networks characterised by large amounts of flux in the number of WSDs, their spectrum needs and traffic, thereby increasing the complexity of the algorithm.

Chen and Huang in [14] designed an evolutionary algorithm to iteratively select the least congested channel among a set of available channels. Anandkumar *et al.* in [15], and Liu and Zhao in [16] used distributed multi-armed bandit learning algorithms for spectrum allocation. However, these models consider both primary and secondary users to be slotted. Li in [17] and Xu *et al.* in [22] use game theoretic solutions for distributed channel selections, but they consider the system model to be static.

Spectrum assignment of white space devices studied in [18–20] have also been modelled as non-cooperative games, but they require complete network information for making the spectrum assignments. Chen and Huang in [21] proposed a distributed learning algorithm for channel selection based on channel data rate, but the solution assumes a fixed channel selection profile of users.

Wicke *et al.* in [23] proposed an approach to competitive multi-agent task allocation in a different setting inspired by bounty hunters. However, in this model, the tasks that agents compete for, are not exclusive and an agent can take another unfinished task by a previous agent and obtain a higher payoff.

In this paper, we propose a model that has a distributed decision making by autonomous agents which co-ordinate with each other using a central shared memory located at the Master. This ensures an efficient method of spectrum management with high spectrum utilisation. Our model considers the network structure to be dynamic as is the case in white spaces and doesn't need complete network information for the operation. With the exception of a centralised model in [13], none of the above papers focus primarily on the white space spectrum characteristics like duration of usage, maximum allowed transmission power and white space devices properties like the frequency range, spectrum demand, transmission capacity, etc. for spectrum assignment. By taking these properties into account, our model provides more value to the white space devices and helps to increase overall network utility and social welfare.

Unlike most of the spectrum assignment models which have a fixed algorithm for spectrum assignment, our model provides different algorithms to choose a channel in the form of strategies. This makes it a rich and flexible model and any new algorithms can be easily integrated with the model by expanding the strategy set of the white space devices.

Each white space network may change overtime and have different network characteristics and constraints at different times. Our model is flexible and adaptable to this due to it being autonomous and self evolving. When the model is deployed on different networks, each network will have a different demographic profile of strategies and different dominant strategies according to the network dynamics. Thus, our model can be trained for different environments and demand patterns and allow it to seamlessly adapt its strategy distribution with changes in load patterns. The demographic profile and demand pattern published by the Master also helps the new WSD entering the system to automatically have the wisdom of the network.

2 System Model

We consider the spectrum assignment model of white space network, at a given location and associated with a single Master, as a multi-agent system where each WSD acts as an autonomous agent. Formally, we define the system as follows:

$$S = (C, A, \psi). \tag{1}$$

Here, C is a set of channels licensed to some primary users, that have allowed WSDs to operate. Each channel $c_i \in C$ is said to be in one of the three different states:

$$state(c_i) = \begin{cases} 0, & \text{if free} \\ 1, & \text{if occupied by primary user} \\ 2, & \text{if occupied by secondary user or WSD.} \end{cases} \tag{2}$$

$A = \{m, W\}$ is a set of autonomous agents where m is the Master WSD, and W is a set of secondary WSDs. The term ψ refers to a set of "strategies" that the system is endowed with. All WSDs have a copy of the set of strategies listed in ψ. Each strategy $\psi_i \in \psi$ denotes a heuristic with which, a WSD makes a choice regarding its requirements.

Demand and Offer. Each WSD has its own set of requirements and constraints, regarding the spectrum. We call this as the "Demand" from the WSD and represent it as a vector: $D = (r_{fr}, r_{Tx}, r_{nc}, r_d)$. Here r_{fr} is the operating frequency range, r_{Tx} is the maximum transmission power of WSD, r_{nc} and r_d are the required bandwidth and duration respectively.

Whenever a WSD wants to communicate in a white space spectrum, it requests the Master for a list of free spectrum fragments. Master obtains information from the geo-location database and forwards it as a set of "Offers" $R = \{O_1, O_2, \ldots, O_k\}$ to the WSD. Each offer $O_j \in R$ is a vector of the form $O_j = (o_{fr}, o_{Tx}, o_{nc}, o_d)$, such that each element of the offer vector *covers* the corresponding element in the demand vector.

While every offer made by the Master is a possible allocation for the WSDs requirements, not all offers bring the same value. Choosing some offers may result in wasted time or bandwidth, while choosing some other offer may have repercussions on choices available to other WSDs.

A WSD does not have the information or resources to compute the impact of its decision on others. It chooses an offer based on local considerations and this choice is defined by the current strategy chosen by the WSD.

Strategy and Fitness. As mentioned earlier, each WSD is endowed with a set ψ of strategies. Initially, a WSD chooses a strategy $\psi_i \in \psi$ uniformly at random. For each of the offers received by the WSD, it computes a gain along several dimensions, namely: feasibility, transmission power, bandwidth, duration and continuity of use. The gains are then tempered differently based on the current strategy, to compute the overall fitness of an offer given the demand.

The feasibility of an offer is a binary variable used to filter away incorrect offers if any. This is computed as follows:

$$f_f = \begin{cases} 1, & \text{if } r_{fr} \text{ is within } o_{fr} \text{ and } o_{nc} \geq r_{nc} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The gains with respect to bandwidth, power and duration requirements are computed as follows:

$$f_n = 1 - \frac{|o_{nc} - r_{nc}|}{max(o_{nc}, r_{nc})} \tag{4}$$

$$f_t = 1 - \frac{|o_{Tx} - r_{Tx}|}{max(o_{Tx}, r_{Tx})} \tag{5}$$

$$f_d = 1 - \frac{|o_d - r_d|}{max(o_d, r_d)}. \tag{6}$$

If a given channel in the offer has been used in the previous transmission of the WSD, it is also said to constitute a gain as it reduces the need to re-calibrate the transmission on a different channel. This is formally represented by a value f_h for a given offer O_j which is set to 1 if the same channel offered by O_j was used in the previous transmission as well, or 0 otherwise.

$$f_h = \begin{cases} 1, & \text{if same channel was used in the previous request} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

A strategy tempers the importance of each of the above gains in different ways, and is represented by a four dimensional simplex $\psi_i = (i_n, i_d, i_t, i_h)$ such that $i_n + i_d + i_t + i_h = 1$.

The fitness of a given offer O_j according to the current strategy ψ_i is given by:

$$F(\psi_i, D, O_j) = f_f \Big(i_n f_n + i_d f_d + i_t f_t + i_h f_h \Big). \tag{8}$$

Once the fitness value is calculated for every offer by comparing it against the demand of the WSD, the offer that has the maximum fitness value is chosen.

$$\max_{\forall O_j \in R} \Big(\; F(\psi_i, D, O_j) \Big). \tag{9}$$

A Representative Bag of Strategies. In our work, the system is endowed with a set of strategies that are explained here. Different strategies compute the fitness of the offer calculated using Eq. (8) in different ways by setting the four dimensional simplex $\psi_i = (i_n, i_d, i_t, i_h)$ of importance scores according to the strategy. The strategy set ψ can be extended by adding more strategies tailored to suit the network requirements or WSD specific strategy.

Longest duration: WSD chooses an offer that has the longest duration o_d. For this strategy, the importance score i_d is very high, whereas i_n, i_t and i_h are low.

Highest allowed transmission power: WSD chooses offer that has the highest permissible transmission power o_{Tx}. Hence, the importance score of i_t is very high and rest of the importance scores are low.

Frequently used: WSD chooses an offer which is frequently used, the importance score of i_h is a high value and rest of the importance scores are low.

Most recently used: WSD chooses an offer that is most recently used. This strategy means WSD wants to continue using the channel it is already using, hence the importance score of i_h is very high (close to 1) and rest of the importance scores are very low (close to 0).

First fit: WSD chooses the first feasible offer. The importance scores for i_d, i_n, i_t are thus assigned 0 and i_h is assigned 1.

Least-biased match: WSD chooses a channel(s) which is close to its operating parameters and to its requirements. This strategy may seem to be best for the total network utilisation, and is also called the "Best match" strategy. The importance scores i_d, i_n, i_t are given equal value, $i_d = i_n = i_t = 1/3$ and $i_h = 0$.

Channel Allocation. WSDs allocate channels for themselves based on the fitness of an offer computed according to the current strategy. They then stake claim on this channel by issuing an atomic TEST&SET operation on the WSDB (via the Master).

TEST&SET tests the state of a channel and allocates it, if the channel is free. This is done in one atomic step to prevent race conditions. If the TEST&SET succeeds, the WSD proceeds to use the channel that it staked claim to. If the operation fails, then the WSD tries to allocate itself to the next best offer based on the fitness value computed.

The actual payoff obtained by a WSD for a given strategy is defined as the fitness value of the offer, which results in a successful allocation of the channel. The payoff obtained at the i^{th} attempt is denoted by ϕ_i. The payoff is set to 0 for every failed attempt.

Demographic Profile. "Demographic profile" is the means by which the WSDs interact with each other. WSDs take advantage of the white space network's infrastructure requirement of communicating with the WSDB via the Master, and piggy back the information about their payoffs to the Master. The

Master uses this information to create the demographic profile, which is the probability distribution of the success of the strategies. Master thus acts as a broker or facilitator among the WSDs without any decision making capabilities. In the absence of such a broker, there would be a need for lot of message exchanges between WSDs which would increase communication complexity.

The initial choice of a strategy by a WSD is made using a uniformly random function. However, over time, as payoffs accumulate in a differentiated fashion across strategies, the system allows for WSDs to change their strategy using an evolutionary rationale.

Following is the calculation used to find the success of a strategy. For each WSD w_k, success of its currently adopted strategy ψ_i is calculated by computing the ratio of sum of positive payoff each WSD gets in comparison with the total number of attempts t_k the WSD makes to get the channel.

$$success(\psi_i, w_k) = \frac{\sum_{i=0}^{t_k} \phi_i}{t_k}. \tag{10}$$

Let $wsds(\psi_i)$ denote the set of all WSDs which have chosen strategy ψ_i. The *demographic dividend* for strategy ψ_i is given by:

$$dividend(\psi_i) = \sum_{\forall w_k \in wsds(\psi_i)} success(\psi_i, w_k). \tag{11}$$

Finally, the *demographic profile* across the system is computed by assigning new probabilities to the strategies using the strategy success calculated above.

$$P(\psi_i) = \frac{dividend(\psi_i)}{\sum\limits_{\forall \psi_i \in \psi} dividend(\psi_i)}. \tag{12}$$

The demographic profile is published by the Master and is continuously updated.

WSD uses a strategy it selected for a certain time period t_p called an *epoch*. At the end of each epoch, the WSD uses the demographic profile published by the Master and optionally changes its strategy. This is controlled by a tuning parameter $\beta \in [0, 1]$, where with a probability β, the WSD continues with its current strategy and with a probability $(1 - \beta)$, it chooses a new strategy, with a random distribution based on the demographic profile.

WSD sets the tuning parameter β based on the history of its payoff. If a WSD is continuously getting higher payoff from the current strategy, it may incrementally increase the value of β. For example, a mobile WSD may set β close to 1 for a strategy that chooses a channel with maximum transmission power, which would allow the WSD to operate seamlessly on the same channel (under the same Master) even if it moves far away from the Master.

WSD may also use a strategy temporarily, like in case of a primary user unexpectedly wanting to transmit on a spectrum which a WSD is already transmitting on. WSD being a secondary device, gives up the spectrum it is currently using and quickly acquires a new channel using a strategy like the "First fit" strategy (picking up feasible offer encountered) which doesn't need much processing. Even though WSD uses a different strategy in such a case, it doesn't shift to this strategy.

System Equilibrium and Training. As WSDs change their strategies, the global demographic profile changes too. We say that the system has reached the state of equilibrium (or, in a state of evolutionary best-response) when the maximum change in the probability for any strategy, falls below a threshold ϵ.

A given deployment of WSDs is "trained" under different "demand profiles" using a simulation framework, until the demographic profile is learned for that demand profile.

We consider three different demand profiles - Peak, Lean and Bursty which vary in the number of WSDs making requests and spectrum demands of the WSD. In Peak Demand Profile (PDP), there are many active WSDs and most WSDs make greater demands on the spectrum. In Lean Demand Profile (LDP), number of WSDs requesting the spectrum is less and most of the WSDs would make low or moderate demands on the spectrum. In Bursty Demand Profile (BDP), WSDs becoming active and demand requirement is random. For each of the demand profiles, a separate demographic profile is published for the consumption of WSD and also each WSD may follow a different strategy for each of the demand profiles.

The Master node can determine the current demand profile based on the traffic that is being routed through it. Based on the training, it also determines the suitable demographic profile for such a demand profile. The Master publishes both these data in the shared memory during deployment.

If the WSD moves from one Master to other, it has to connect with the new Master, however, it will be able to quickly adapt to the new network and choose a strategy and channel using the demographic profile of the new Master.

3 Evaluation

We conducted experimental simulation to quantify the performance of our model using an open-source simulation package $OMNET{++}$. For our experiments, we used the data from the study conducted by one of our partners IIM-Ahmedabad, India [24], on internet usage data in rural India. We also used the data from Google spectrum database[1].

We considered 20 free channels in our experiments. We modelled the three different demand profiles as follows: In PDP, WSDs becoming active is modelled as a Poisson distribution with a value of $\lambda \geq 15$ active WSDs/hour, in LDP, it is modelled as a Poisson distribution with a value of $\lambda \leq 8$ active WSDs/hour and in BDP, it is modelled as a Gaussian distribution. Primary users being active was also modelled according to the data mentioned above.

[1] http://www.google.com/get/spectrumdatabase/.

Demand Cards. WSD's spectrum requirements are simulated in the form of *Demand Cards* which are representative of the combinations of the usage parameters of the WSD. Each demand profile is associated with a deck of *Demand Cards* with different probabilities assigned to them. A WSD chooses a card from the current deck uniformly at random, to simulate its demand.

Table 1. Demand Cards

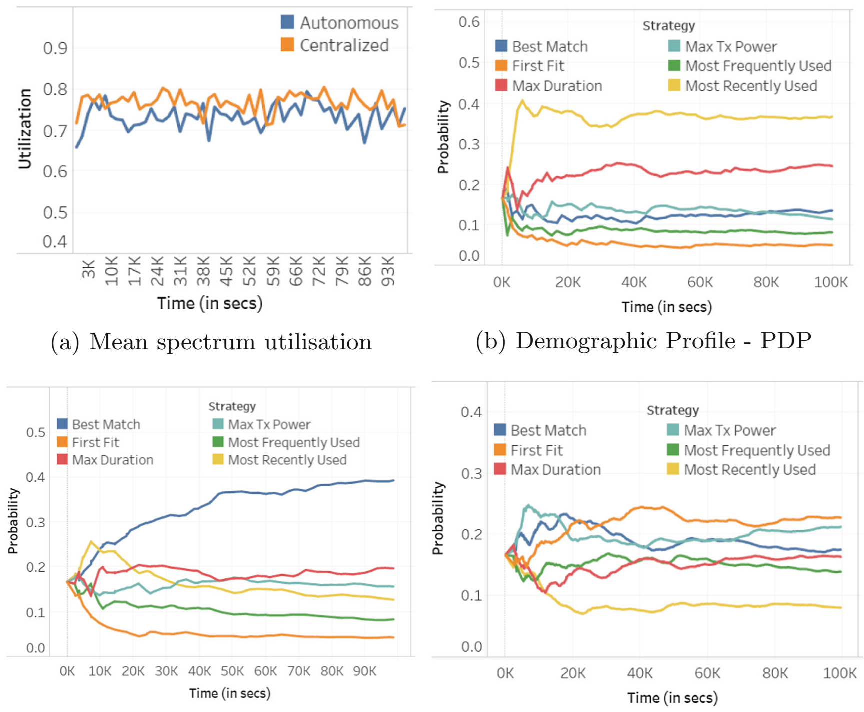
Card	No of channels	Duration	Transmission power
Card 1	1	1–3 h	23 dBM
Card 2	2	3–5 h	26 dBM
Card 3	2	3–5 h	36 dBM
Card 4	4	5–8 h	30 dBM
Card 5	3	>8 h	36 dBM

Table 1 shows few of the demand cards that we used for our simulations (Out of 60+ demand cards created with this data). These demand cards were created based on the study on internet usage data in rural India, which included data on average number of hours spent per day by the user in accessing internet, internet usage patterns on different days, type of applications or services used, type of content used (video, audio streaming, web conferencing etc.) and their usage percentage.

For different demand profiles, different cards are chosen by the WSDs. During PDP, WSD chooses cards like 4 and 5, that make higher demands on spectrum with higher probability, than other cards. During LDP, WSD chooses cards like 1 having less spectrum demands with higher probability, and 2 and 3 with moderate probability. During BDP, WSD chooses any of the cards from the deck with equal probability. Similarly, we created demand cards of different combinations of the above parameters for the data obtained from Google spectrum database and used them in our simulations.

Evaluation Results. We evaluated the efficiency of our model by comparing the spectrum allocation of the white space network done by our autonomous spectrum assignment model with the spectrum allocation done by a centralised model. We considered centralised model of spectrum allocation as the benchmark, since it represents the ideal allocation for all the WSDs due to its global knowledge and the ability to freeze network operations while performing the allocation.

Figure 1a shows the mean spectrum utilisation for both the models for PDP. We can see that the utilisation for autonomous model is comparable to that of the utilisation for the centralised model. We found similar results for LDP and BDP demand profiles. Figure 1b, c and d show graphs of probability of usage of each strategy at different time intervals for PDP, LDP and BDP demand profiles



(a) Mean spectrum utilisation
(b) Demographic Profile - PDP
(c) Demographic Profile - LDP
(d) Demographic Profile - BDP

Fig. 1. Mean spectrum utilisation and demographic profiles

respectively. At each time instance, the probability values of all the strategies constitute the demographic profile at that instance. We can see that one of the strategies emerges with high probability in all the three demand profiles. For example, in PDP (Fig. 1b), "Most recently used" strategy has higher probability than other strategies, followed by "Maximum duration". We can also see that the demographic profile stabilises after sometime and the probabilities of strategies do not change beyond a threshold ϵ (0.01 in this case).

We ran simulations for different combinations of WSDs and different values of network characteristics. In all the cases, we observed that the strategies stabilise overtime and the system reaches a state of equilibrium. This also indicates that none of the WSDs are starving. We can conclude this because, if the WSDs were in a state of starvation, they would have changed their strategies in order to get better payoff and hence the demographic profile would not have stabilised. To further confirm the fairness of our model, we compared the mean of average payoffs of all the WSDs in our model with the centralised model and found them to be close.

Our algorithms that form various strategies are efficient and the fitness function which finds an optimal channel from all the available feasible channel(s) runs in linear time.

4 Conclusions

In this paper, an autonomous-agent model for spectrum assignment of white space devices is presented. The self-evolving and autonomous white space devices allocate spectrum to themselves in a way that maximises their utility, while also maximising the overall spectrum utility of the network. The developed solution of autonomous spectrum assignment can be deployed and self-trained for any geographical area with different internet usage patterns and network topology for optimum network utilisation. It is also flexible to adapt its strategy distribution with changes in load patterns or network dynamics. The model not only provides different algorithms to choose a channel in the form of strategies, but can also be customised and enriched with new algorithms. In our future work, we plan to expand the strategy set for the white space devices. We also propose the WSDs to have heterogeneous strategy set according to their idiosyncrasies, as against the current WSDs which have the same set of strategies. We also plan to include topological parameters like path loss and terrain information into our model for optimising network utilisation.

Acknowledgements. This work is partially supported by EU-India REACH Project under Grant ICI+/2014/342-896. The project aims to develop advanced technical solutions for providing high-speed broadband internet access in rural India in the unlicensed white space spectrum.

References

1. Zhu, L., Chen, V., Malyar, J., Das, S., McCann, P.: Protocol to access white-space (PAWS) databases (2015)
2. Murty, R., Chandra, R., Moscibroda, T., Bahl, P.: Senseless: a database-driven white spaces network. IEEE Trans. Mob. Comput. **11**(2), 189–203 (2012)
3. Li, C., Liu, W., Li, J., Liu, Q., Li, C.: Aggregation based spectrum allocation in cognitive radio networks. In: 2013 IEEE/CIC International Conference on Communications in China-Workshops (CIC/ICCC), pp. 50–54. IEEE (2013)
4. Halldórsson, M.M., Halpern, J.Y., Li, L.E., Mirrokni, V.S.: On spectrum sharing games. In: Proceedings of the Twenty-Third Annual ACM Symposium on Principles of Distributed Computing, pp. 107–114. ACM (2004)
5. Cao, L., Zheng, H.: Distributed spectrum allocation via local bargaining. In: SECON, pp. 475–486 (2005)
6. Chen, D., Zhang, Q., Jia, W.: Aggregation aware spectrum assignment in cognitive ad-hoc networks. In: 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom 2008, pp. 1–6. IEEE (2008)
7. Cao, L., Zheng, H.: Distributed rule-regulated spectrum sharing. IEEE J. Sel. Areas Commun. **26**(1), 130–145 (2008)

8. Nie, N., Comaniciu, C.: Adaptive channel allocation spectrum etiquette for cognitive radio networks. Mob. Netw. Appl. **11**(6), 779–797 (2006)
9. Suris, J.E., DaSilva, L.A., Han, Z., MacKenzie, A.B.: Cooperative game theory for distributed spectrum sharing. In: IEEE International Conference on Communications, ICC 2007, pp. 5282–5287. IEEE (2007)
10. Bourdena, A., Kormentzas, G., Pallis, E., Mastorakis, G.: A centralised broker-based CR network architecture for TVWS exploitation under the rtssm policy. In: 2012 IEEE International Conference on Communications (ICC), pp. 5685–5689. IEEE (2012)
11. Bourdena, A., Kormentzas, G., Skianis, C., Pallis, E., Mastorakis, G.: Real-time TVWS trading based on a centralized CR network architecture. In: 2011 IEEE GLOBECOM Workshops (GC Workshop), pp. 964–969. IEEE (2011)
12. Pei, Y., Ma, Y., Peh, E.C.Y., Oh, S.W., Tao, M.-H.: Dynamic spectrum assignment for white space devices with dynamic and heterogeneous bandwidth requirements. In: 2015 IEEE Wireless Communications and Networking Conference (WCNC), pp. 36–40. IEEE (2015)
13. Kash, I.A., Murty, R., Parkes, D.C.: Enabling spectrum sharing in secondary market auctions. IEEE Trans. Mob. Comput. **13**(3), 556–568 (2014)
14. Chen, X., Huang, J.: Evolutionarily stable open spectrum access in a many-users regime. In: 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011), pp. 1–5. IEEE (2011)
15. Anandkumar, A., Michael, N., Tang, A.: Opportunistic spectrum access with multiple users: learning under competition. In: 2010 Proceedings IEEE INFOCOM, pp. 1–9. IEEE (2010)
16. Liu, K., Zhao, Q.: Decentralized multi-armed bandit with multiple distributed players. In: Information Theory and Applications Workshop (ITA), pp. 1–10. IEEE (2010)
17. Li, H.: Multi-agent q-learning of channel selection in multi-user cognitive radio systems: a two by two case. In: IEEE International Conference on Systems, Man and Cybernetics, SMC 2009, pp. 1893–1898. IEEE (2009)
18. Felegyhazi, M., Čagalj, M., Hubaux, J.-P.: Efficient MAC in cognitive radio systems: a game-theoretic approach. IEEE Trans. Wireless Commun. **8**(4), 1984–1995 (2009)
19. Niyato, D., Hossain, E.: Competitive spectrum sharing in cognitive radio networks: a dynamic game approach. IEEE Trans. Wireless Commun. **7**(7), 2651–2660 (2008)
20. Han, Z., Pandana, C., Liu, K.: Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning. In: IEEE Wireless Communications and Networking Conference, WCNC 2007, pp. 11–15. IEEE (2007)
21. Chen, X., Huang, J.: Spatial spectrum access game: nash equilibria and distributed learning. In: Proceedings of the Thirteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, pp. 205–214. ACM (2012)
22. Xu, Y., Wang, J., Wu, Q., Anpalagan, A., Yao, Y.-D.: Opportunistic spectrum access in cognitive radio networks: global optimization using local interaction games. IEEE J. Sel. Top. Sig. Process. **6**(2), 180–194 (2012)
23. Wicke, D., Freelan, D., Luke, S.: Bounty hunters and multiagent task allocation. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, pp. 387–394. International Foundation for Autonomous Agents and Multiagent Systems (2015)
24. Jain, R.: REACH Internal Technical Report - Internet usage data in rural India. IIM Ahmedabad, India (2017)